Complete ^{13}C assignments for recombinant Cu(I) rusticyanin Prediction of secondary structure from patterns of chemical shifts

Anna Toy-Palmer**, Stefan Prytulla, H. Jane Dyson*

The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037, USA

Received 24 March 1995

Abstract Complete resonance assignments for the ^{13}C spectrum of reduced (Cu(I)) rusticyanin have been made using ^{13}C,^{15}N doubly labeled recombinant material. The reported assignments include those for the carboxyl and carbonyl carbon atoms and protonated aromatic ring carbons, and were obtained using a variety of 2- and 3D inverse-detected NMR experiments, including ^{13}C,^{15}N,^{1}H triple resonance experiments and HCCH-COSY and -TOCSY. Backbone carbonyl assignments were obtained using 3D HNCO and HCACO spectra, and modified versions of 2D H(CA)CO and HMBC spectra were used to obtain side-chain carboxyl carbon and methionine ϵ-methyl carbon assignments, respectively. A comparison of the $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and ^{13}CO chemical shifts with published 'random coil' values confirms the conclusion reached from a consideration of the $^{3}J_{HN\alpha}$ coupling constants and the pattern of sequential NOEs, that the protein consists largely of β-structure.

Key words: Rusticyanin; Type 1 copper protein; Secondary structure

1. Introduction

Rusticyanin is a Type I copper protein of unknown structure obtained from the iron-oxidizing bacterium *Thiobacillus ferrooxidans*, which occurs in acid mine drainage [1]. The organism can metabolize both elemental sulfur and metal sulfides; under conditions where the oxidation of Fe^{2+} is the principal source of electrons for metabolism, rusticyanin may constitute up to 5% of the total protein in the cell [2]. One of the by-products of the oxidation of metal sulfides by *Thiobacillus* is sulfuric acid, and the organism shows a preference for pHs lower than 4.0, consistent with the lower rates of Fe^{2+} autoxidation at acid pH. Rusticyanin forms part of the initial electron transfer chain from Fe^{2+} (though it is probably not the initial electron acceptor [3]) and is found in the periplasm of the cell. The pH optimum for stability of rusticyanin is therefore close to the acidic conditions of the external medium, rather than the neutral pH of the interior of the cell. One of the interesting questions about rusticyanin is the source of the stability of the copper site under acidic conditions. For almost all Type 1 copper proteins the copper is labile under these conditions, due to protonation of one or both of the histidine ligands [4]. Unusual copper ligands can apparently be ruled out, since rusticyanin appears to contain the classic Type 1 copper site (His_2 Cys Met) [5,6]. It therefore appears that the structure of the protein and details of the local geometry of the copper mediate the stability of the copper site under acidic conditions. Structural and dynamic information on this protein obtained using NMR measurements can be used to probe this process.

Assignment of the ^{1}H and ^{15}N NMR spectra of reduced (Cu(I)) rusticyanin has been largely completed using ^{15}N-labeled protein obtained from *Thiobacillus* [5]. Excellent yields of protein uniformly labeled with ^{13}C and ^{15}N have been obtained from the over-expression in *Escherichia coli* of a synthetic gene [6]. We here report the assignment of the complete ^{13}C spectrum of rusticyanin, obtained from 3-dimensional heteronuclear and triple-resonance spectra. The data provide an excellent example of the correlation of chemical shifts with secondary structure, as judged from inter-residue NOE connectivities and $^{3}J_{HN\alpha}$ coupling constants.

2. Materials and methods

Recombinant rusticyanin uniformly labeled with ^{13}C and ^{15}N was obtained using M9 minimal medium as described previously [6], with the addition of 2 g/l D-glucose uniformly labeled with ^{13}C in place of the unlabeled glucose previously used. A yield of pure protein of 50 mg/litre of medium was obtained.

The doubly-labeled recombinant protein was purified and prepared for NMR spectroscopy as previously described [6]. The molecular weight of the purified copper-containing protein was 17,557 Da by electrospray mass spectrometry, indicating a level of incorporation of ^{13}C of >99% assuming 100% incorporation of ^{15}N. Samples were routinely used between 2 and 5 mM in 1 mM H_2SO_4, pH 3.4.

NMR spectra were acquired on Bruker spectrometers operating at 500 and 600 MHz for protons. ^{13}C HSQC and constant-time HSQC [7] spectra were obtained using combined States-TPPI quadrature detection, with a spectral width of 12.5 ppm (6,250 Hz) and 1,024 complex points in ω_2 (^{1}H), and 80.5 ppm (10,060 Hz) and 256 complex points in ω_1(^{13}C). A 2D H(CA)CO experiment [8] with spectral widths 10.4 ppm (6,250 Hz) and 2,048 complex points in ω_2 (^{1}H) and 41.4 ppm (6,250 Hz) and 512 complex points in ω_1 (^{13}CO) was acquired using TPPI quadrature detection. An internal reference of dioxane (singlet resonance at 3.75 ppm from DSS) was used for proton chemical shifts, and carbon chemical shifts were indirectly referenced according to the method of Live et al. [9], using a ratio of 0.25145002.

Three-dimensional heteronuclear-correlated spectra of samples dissolved in $^{2}H_2O$ were used to assign the aliphatic ^{13}C resonances of rusticyanin. Both HCCH-COSY [10] and HCCH-TOCSY [11] spectra were utilized, with spectral widths of 12.5 ppm (6,250 Hz,) and 1024 complex points in ω_3 (^{1}H), 7 ppm (3,500 Hz) and 120 complex points in ω_1 (^{1}H) and 40 ppm (5,030 Hz) and 44 (HCCH-COSY) or 50 (HCCH-TOCSY) ^{13}C planes in ω_2. The carbon dimension was folded for maximum digital resolution.

Three-dimensional triple-resonance spectra were acquired, including HBHA(CO)NH [12], CBCA(CO)NH [12,13] and C(CO)NH-TOCSY [14]. Spectral widths for these spectra were 8 ppm (4,000 Hz), with 1024 complex points in ω_3 (^{1}H), 30 ppm (1,520 Hz) with 32 complex points in ω_2 (^{15}N) and 70 ppm (8,803 Hz) and 40 (CBCA(CO)NH) or 51 (C(CO)NH-TOCSY) complex points in ω_1 (^{13}C). The spectral width was 6.5 ppm (3251 Hz) in ω_1, with 64 complex points for the HBHA(CO)NH spectrum. These spectra were acquired using semi-

*Corresponding author.

**Present address:* Tripos Inc., 1699 South Hanley Road, St. Louis, MO 61344, USA.

SSDI 0014-5793(95)00423-8

constant time in ω_1. The backbone carbonyl carbon assignments were made using a 3D HNCO spectrum [15] (constant-time in ω_1 and semi-constant-time in ω_2) acquired with spectral widths 12.5 ppm (6,250 Hz) and 1,024 complex points in ω_3 (^{1}H), 20 ppm (2,514 Hz) and 64 complex points in ω_2 (^{13}CO) and 40 ppm (2,026 Hz) and 24 complex points in ω_1 (^{15}N). The 3D HCACO spectrum [8] was acquired with spectral widths 5.0 ppm (2,500 Hz) and 1,024 complex points in ω_3 (^{1}H), 19.9 ppm (2,500 Hz) and 32 complex points in ω_2 ($^{13}C^{\alpha}$) and 40 ppm (2,500 Hz) and 32 complex points in ω_1 (^{13}CO) A modification of the H(CA)CO experiment was used [16,17] to determine the assignments of the carboxyl carbon atoms of Asp, Glu, Asn and Gln. Assignments of all proton-attached aromatic carbons were made using an HSQC spectrum where the offset frequency had been shifted to optimize for excitation of aromatic carbon and proton frequencies. Data sizes and spectral widths were the same for this spectrum as for all other HSQC spectra (see above), and the ^{13}C carrier frequency was set to 140 ppm. The ε-methyl carbon resonances of the methionine residues were assigned using a modified version of the HMBC experiment [18] optimized for the detection of long-range coupling through sulfur, with spectral widths of 12.5 ppm (6,250 Hz) and 1,024 complex points in ω_2 ($^1H^{\varepsilon}$) and 65 ppm (8,170 Hz) and 64 complex points in ω_1 ($^{13}C^{\gamma}$).

All spectra were Fourier transformed and analyzed using either a modified version of the FTNMR software provided by Dr. Dennis Hare, FELIX 2.3 (Biosym Technologies) or the NMR-TRIAD software from Tripos, Inc. 2D spectra were commonly zero-filled to give 2,048 × 2,048 real matrices, while the 3D spectra generally contained 64 planes of 1,024 × 1,024 real data points. Baseline corrections were applied when necessary, and a low-pass filter was used to remove the residual water resonance in the triple-resonance experiments.

3. Results and discussion

3.1. Resonance assignments

Triple-resonance and HCCH-TOCSY experiments were the most useful in determining aliphatic ^{13}C resonance assignments for rusticyanin, and the process was assisted by the availability of nearly-complete proton and ^{15}N resonance assignments [5].

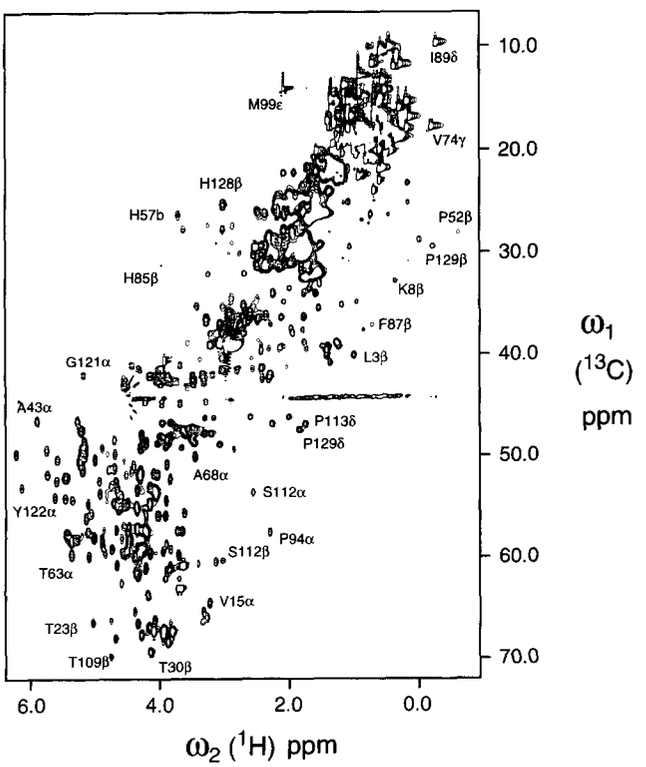

Fig. 1. 500 MHz ^{13}C HSQC spectrum of Cu(I) rusticyanin. Selected resonances have been labeled.

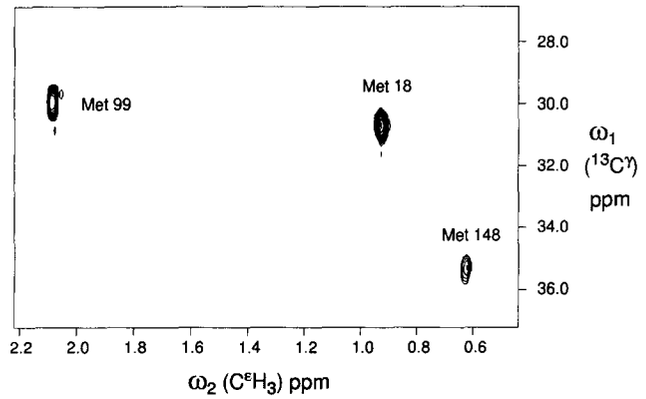

Fig. 2. 500 MHz HMBC spectrum modified to show long-range connectivities between the C^{γ} and $C^{\varepsilon}H_3$ resonances of the three methionine residues in rusticyanin.

The triple-resonance experiments also served to confirm the sequential assignments, which had relied heavily on inter-residue NOEs. For a β-sheet protein such as rusticyanin, there is a risk in using only NOEs to establish sequential connectivities, since the $d_{\alpha N}(i,i+1)$ NOE is almost the only connection between spin systems; overlap in the $C^{\alpha}H$ region can produce unacceptable ambiguities. Since ^{13}C-labeled rusticyanin was not available at that time, the ^{1}H and ^{15}N assignments were confirmed using an ^{15}N relayed-NOE experiment [5]. The present work demonstrates the validity of this approach: all of the sequential assignments made on the basis of the NOE and relayed-NOE data have now been confirmed by triple-resonance spectra.

The dispersion of the ^{13}C-^{1}H HSQC spectrum is excellent, as shown in Fig. 1. The assignment of aliphatic ^{13}C atoms with protons attached was achieved by a combination of information from CBCA(CO)NH and C(CO)NH-TOCSY spectra and from HCCH-TOCSY and HCCH-COSY spectra. The $C^{\varepsilon}H_3$ assignments for the methionine residues were made using a modified HMBC spectrum [18] which directly correlates the methyl protons with the C^{γ} carbon, as shown in Fig. 2; the C^{ε} assignment was then obtained from the ^{13}C-^{1}H HSQC spectrum. Aromatic proton-attached ^{13}C assignments were made from an HSQC spectrum centered on the aromatic region: overlap is quite severe in some parts of this spectrum, due to the large number of Phe residues.

The only ^{13}C assignments made for carbon atoms without directly-bonded protons were those of the backbone carbonyl carbons (made using 3D HNCO and HCACO spectra) and of the carboxyl carbons of Asp and Glu and the carboxamide carbons of Asn and Gln, which were made using a 2D modified H(CA)CO experiment that correlates the carboxyl carbon with the next-door methylene protons, either $C^{\beta}H$ or $C^{\gamma}H$ [16,17]. With the exception of Gly^1, the $C^{\alpha}H$-CO correlations of glycine residues were absent from a normal 2D H(CA)CO spectrum, and several were missing from the 3D HCACO spectrum. However, as shown in Fig. 3, these resonances were present as strong cross peaks in the modified H(CA)CO spectrum used to assign the side chain carboxyl carbons, most likely due to the similarity of the ^{13}C resonance frequencies of the glycine C^{α} and the aspartate/asparagine C^{β}. This is confirmed by the identification of several cross peaks in the spectrum that arise from backbone

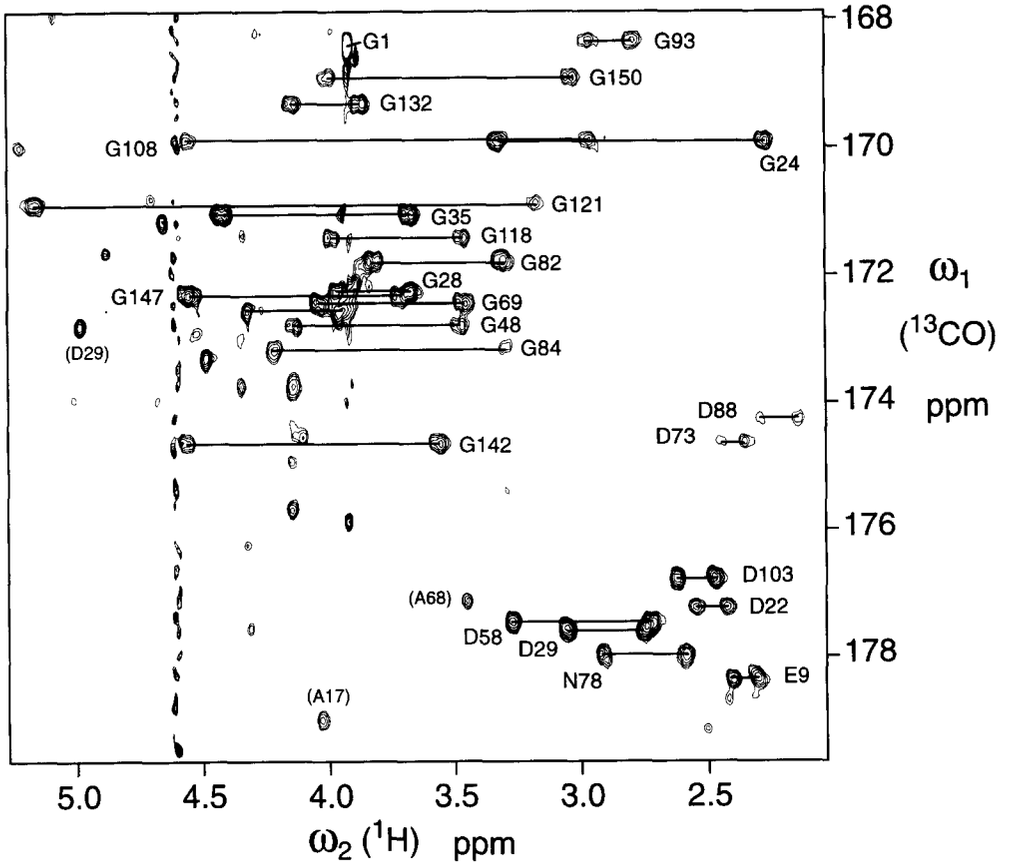

Fig. 3. Portion of the modified 2D H(CA)CO spectrum used to assign side chain carboxyl and carboxamide carbon resonances and to confirm the backbone CO assignments for the glycine residues. Connectivities for the glycine residues and for selected side chains are labeled; the remaining side chain cross peaks are of lower intensity. Cross peaks labeled in parentheses and unlabeled cross peaks arise from other backbone H(CA)CO connectivities where the C^{α} chemical shift falls within the excitation profile for this experiment.

H(CA)CO connectivities for other residues such as Ala[68], Ala[17] and Asp[29] whose C^{α} resonances are at relatively high field. Assignments were made for all but two of the the backbone carbonyl resonances. The backbone CO assignments for the residues preceding the 14 prolines were not available from the HNCO spectrum, but could in most cases be obtained unambiguously from the 3D HCACO spectrum. The two exceptions are Leu[12] and His[128]. The ^{13}C assignments for all residues are shown in Table 1.

3.2. ^1H assignments for proline residues

Although all of the backbone resonances with attached protons had been assigned previously using ^{15}N spectra [5], a number of ^{1}H resonances remained unassigned, including those of the proline residues, which are not usually accessible in ^{15}N spectra because they have no amide proton. Complete ^{13}C assignments were obtained from the CBCA(CO)NH and C(CO)NH-TOCSY spectra for all of the proline residues in rusticyanin except for Pro[94], which is followed by another proline, and ^{1}H assignments for the $C^{\alpha}H$ and $C^{\beta}H$ were obtained unambiguously from the HBHA(CO)NH experiment. The remaining resonance assignments were completed using the HSQC and HCCH-TOCSY experiments. A summary of the ^{1}H assignments of the proline residues is shown in Table 2. It is noteworthy that several of the ^{1}H resonances are highly upfield-shifted; this is probably due to ring-current effects from tightly-packed aromatic residues close to the copper site. These observations are probably significant with respect to the relationship between structure and function in rusticyanin. Several of the proline residues show evidence of a *cis* configuration of the X-Pro peptide bond. For example, the Phe[51]-Pro[52] peptide bond is most likely *cis*, shown not only by the values of the ^{13}C chemical shifts for Pro[52] (Table 1) but also by the presence of an NOE between the $C^{\alpha}H$ protons of the two residues. At present, we can define 6 prolines as *trans* on the basis of $d_{\alpha\delta}(i,i+1)$ NOEs (residues 47, 50, 67, 100, 113 and 129), two others as likely to be *trans* (residues 94 and 141) and two as definitely *cis* on the basis of $d_{\alpha\alpha}(i,i+1)$ NOEs (residues 52 and 95). The isomerization states of the remaining three residues are still undefined, due to resonance overlaps and lack of visible NOE connectivities. It is interesting to note that the two *cis*-prolines are preceded by residues, Phe and Pro, that are known to have a propensity for the formation of *cis*-proline in unstructured peptides in solution [19,20], and further that other Phe-Pro sequences are among those that are *trans*. The formation of *cis* peptide bonds in proteins appears to be highly dependent on the protein structure and only secondarily on the actual sequence.

Table 1
^{13}C resonance assignments for rusticyanin

Residue	CO	C^{α}	C^{β}	C^{γ}	C^{δ}	Other
Gly[1]	168.4	40.7	–			
Thr[2]	172.6	59.3	67.3	19.4		
Leu[3]	174.5	51.7	40.1	24.4	20.2(0.43),21.9(0.56)	
Asp[4]	173.8	49.9	34.6	174.2		
Thr[5]	171.4	57.3	66.2	18.2		
Thr[6]	171.1	59.0	67.3	19.3		
Trp[7]	173.8	53.6	29.1			120.1(C^{δ}),118.0(C^{ε}),119.6($C^{\zeta 3}$), 121.2(C^{η}),113.2($C^{\zeta 2}$)
Lys[8]	172.7	51.9	32.9	23.3	26.9	40.0
Glu[9]	174.2	52.1	29.6	30.1	177.9	
Ala[10]	175.4	47.5	21.3			
Thr[11]	174.0	57.9	68.0	19.5		
Leu[12]		57.7	36.4	24.7	22.7(0.95),19.5(0.71)	
Pro[13]	178.4	63.7	28.5	26.2	46.7	
Gln[14]	177.2	55.8	26.0	31.7	176.9	
Val[15]	175.5	64.5	29.2	17.8(0.49),21.2(0.70)		
Lys[16]	177.0	57.4	29.5	23.0	26.9	39.4
Ala[17]	179.0	51.9	15.2			
Met[18]	177.3	57.4	31.6	30.3		12.9
Leu[19]	178.2	55.6	38.9	24.1	20.1(0.60),24.0(0.70)	
Glu[20]	174.2	53.6	25.9	30.5	178.8	
Lys[21]	173.6	53.0	29.0	22.4	26.3	39.7
Asp[22]	176.6	52.5	41.5	176.9		
Thr[23]	173.5	59.6	66.4	20.2		
Gly[24]	169.7	42.1				
Lys[25]	174.4	51.5	32.3	22.5	26.6	39.7
Val[26]	174.6	59.9	30.7	19.3(0.91),19.9(0.80)		
Ser[27]	173.3	54.6	60.9			
Gly[28]	172.2	44.9				
Asp[29]	172.4	50.3	37.2	177.2		
Thr[30]	171.2	58.4	69.3	19.4		
Val[31]	172.6	59.1	30.4	18.9(0.35),17.7(0.64)		
Thr[32]	172.0	59.6	67.2	19.2		
Tyr[33]	172.0	55.0	37.8			131.1(C^{δ}),116.3(C^{ε})
Ser[34]	170.9	54.7	63.1			
Gly[35]	170.9	41.1				
Lys[36]	175.6	55.3	30.7	22.7	26.3	39.7
Thr[37]	171.8	58.2	67.8	19.1		
Val[38]	170.8	57.1	32.4	19.7(0.49),18.4(0.52)		
His[39]	170.4	49.5	27.8			117.7(C^{δ}),134.1(C^{ε})
Val[40]	173.8	58.2	31.7	16.8(0.58),16.9(0.15)		
Val[41]	171.9	59.0	30.1	16.0(0.31),19.0(0.60)		
Ala[42]	172.9	46.5	21.1			
Ala[43]	174.9	46.5	20.4			
Ala[44]	171.7	48.4	17.9			
Val[45]	172.8	60.4	26.5	19.0(0.84),20.6(0.78)		
Leu[46]	174.1	51.0	39.2	24.7	23.0(0.61),20.8(0.71)	
Pro[47]	175.9	61.8	29.4	25.5	47.3	
Gly[48]	172.6	42.2				
Phe[49]	172.1	54.2	35.2			130.9(C^{δ}),127.6(C^{ε}),127.7(C^{ζ})
Pro[50]	174.1	59.9	30.2	25.3	48.1	
Phe[51]	171.8	51.1	38.8			130.2(C^{δ}),129.7(C^{ε}),127.9(C^{ζ})
Pro[52]	171.4	60.6	28.0	20.8	46.6	
Ser[53]	168.7	56.1	62.2			
Phe[54]	173.2	54.1	40.4			129.1(C^{δ}),132.0(C^{ε}),128.6(C^{ζ})
Glu[55]	174.3	52.5	30.7	30.8		
Val[56]	173.8	59.0	32.5	17.0(0.70),19.8(1.01)		
His[57]	172.6	54.4	26.4			117.5(C^{δ}),133.9(C^{ε})
Asp[58]	171.0	52.3	36.6	177.1		
Lys[59]	173.6	49.8	30.7	20.9	25.2	39.4
Lys[60]	173.4	53.5	30.6	22.2	26.9	39.1
Asn[61]	170.3	51.1	36.2	175.1		
Pro[62]	176.2	60.7	29.6	25.5	47.5	
Thr[63]	171.8	59.8	66.8	19.8		
Leu[64]	174.1	50.4	39.3	24.7	20.6(0.62),21.9(21.9)	
Glu[65]	172.7	52.1	25.9	29.6	178.0	
Ile[66]	170.7	54.7	40.9	22.5	10.1	14.3($C^{\gamma}H_3$)
Pro[67]	174.6	59.6	29.7	25.2	48.9	
Ala[68]	177.2	50.1	15.4			

Table 1 (*continued*)

Residue	CO	C^{α}	C^{β}	C^{γ}	C^{δ}	Other
Gly[69]	172.3	42.7				
Ala[70]	175.6	49.3	16.8			
Thr[71]	170.9	61.0	66.1	19.2		
Val[72]	172.2	57.6	30.1	18.7(0.53),19.1(0.35)		
Asp[73]	172.2	49.3	35.9	174.3		
Val[74]	173.4	58.5	28.9	17.9(−0.18),18.0(0.19)		
Thr[75]	170.7	58.2	68.4	19.8		
Phe[76]	171.5	51.8	39.2			129.0(C^{δ}),132.2(C^{ε}),130.8(C^{ζ})
Ile[77]	172.7	56.8	39.4	26.5	13.5	16.7($C^{\gamma}H_3$)
Asn[78]	173.8	49.4	36.1	177.7		
Thr[79]	171.4	58.1	65.4	19.3		
Asn[80]	170.7	49.7	39.4	174.4		
Lys[81]	176.0	55.9	30.1	23.0	27.1	39.0
Gly[82]	171.6	42.8				
Phe[83]	172.5	54.9	41.3			130.0(C^{δ}),130.3(C^{ε}),129.1(C^{ζ})
Gly[84]	173.1	44.7				
His[85]	174.1	54.3	31.3			119.5(C^{δ}),135.5(C^{ε})
Ser[86]	167.7	54.6	66.4			
Phe[87]	173.1	55.3	37.1			128.8(C^{δ}),128.5(C^{ε}),125.3(C^{ζ})
Asp[88]	169.7	50.4	39.4	174.0		
Ile[89]	172.4	57.9	35.2	25.3	9.8	15.4($C^{\gamma}H_3$)
Thr[90]	170.5	56.1	67.5	17.7		
Lys[91]	175.7	55.2	30.7	24.1	26.8	38.7
Lys[92]	173.7	53.6	29.6	21.1	25.8	39.7
Gly[93]	168.1	42.2				
Pro[94]	172.4	57.5	28.1	22.5	46.2	
Pro[95]	174.0	58.6	31.6	22.1	47.8	
Tyr[96]	173.0	55.0	38.0			130.2(C^{δ}),116.8(C^{ε})
Ala[97]	175.4	48.0	17.3			
Val[98]	173.9	65.4	28.4	18.5(0.88),20.2(1.00)		
Met[99]	171.9	48.6	28.3	29.9		14.1
Pro[100]	174.3	61.0	29.9	24.4	48.1	
Val[101]	174.5	59.6	27.9	18.0(0.94),18.2(0.93)		
Ile[102]	173.3	57.2	37.5	21.9	12.1	16.3($C^{\gamma}H_3$)
Asp[103]	172.9	50.3	36.5	176.5		
Pro[104]	173.5	59.3	31.4	22.5	47.9	
Ile[105]	175.2	59.3	35.5	25.3	10.6	15.4($C^{\gamma}H_3$)
Val[106]	174.1	61.4	29.5	19.1(1.09),18.2(1.03)		
Ala[107]	172.3	48.7	18.6			
Gly[108]	169.7	42.4				
Thr[109]	176.6	57.7	69.8	18.6		
Gly[110]	172.4	41.4				
Phe[111]	174.5	57.3	35.4			129.7(C^{δ}),129.5(C^{ε}),127.5(C^{ζ})
Ser[112]	169.2	53.6	60.3			
Pro[113]	175.6	59.3	29.4	24.9	46.9	
Val[114]	172.7	57.1	28.3	16.3(0.69),18.4(0.66)		
Pro[115]	174.6	61.5	29.4	24.9	48.9	
Lys[116]	174.2	53.3	32.9	21.9	26.9	39.4
Asp[117]	173.0	51.8	35.2	174.1		
Gly[118]	171.2	42.9				
Lys[119]	173.4	51.4	32.4	22.7	26.6	40.0
Phe[120]	173.6	54.0	38.7			129.6(C^{δ}),127.7(C^{ε}),127.5(C^{ζ})
Gly[121]	170.7	42.0				
Tyr[122]	170.7	53.1	41.5			131.6(C^{δ}),116.4(C^{ε})
Thr[123]	168.2	57.0	65.9	17.1		
Asp[124]	172.4	49.7	42.4	174.3		
Phe[125]	170.6	53.7	37.9			130.8(C^{δ}),127.9(C^{ε}),125.7(C^{ζ})
Thr[126]	171.6	58.3	68.3	19.4		
Trp[127]	171.3	51.4	30.1			123.9(C^{δ}),116.4(C^{ε}),118.4($C^{\zeta 3}$), 122.1(C^{η}),112.8($C^{\zeta 2}$)
His[128]		49.7	25.4			116.5(C^{δ}),133.5(C^{ε})
Pro[129]	175.7	59.7	29.5	23.0	47.5	
Thr[130]	172.2	57.1	67.6	19.3		
Ala[131]	175.4	50.8	16.1			
Gly[132]	169.2	42.2				
Thr[133]	171.6	59.9	67.5	19.2		
Tyr[134]	169.8	52.4	40.4			132.1(C^{δ}),115.3(C^{ε})
Tyr[135]	172.7	54.3	39.7			130.8(C^{δ}),115.0(C^{ε})
Tyr[136]	172.5	50.5	37.6			130.9(C^{δ}),116.4(C^{ε})
Val[137]	171.9	57.2	34.1	18.7(0.32),18.5(0.39)		

Table 1 (*continued*)

Residue	CO	C^{α}	C^{β}	C^{γ}	C^{δ}	Other
Cys^{138}	175.6	55.6	32.1			
Gln^{139}	174.1	54.7	30.1	34.0	177.5	
Ile^{140}	174.3	59.2	33.2	26.8	10.6	11.8($C^{\gamma}H_3$)
Pro^{141}	175.1	62.5	29.0	25.2	48.8	
Gly^{142}	174.5	43.3				
His^{143}	175.7	57.4	27.3			114.1(C^{δ}),134.8(C^{ε})
Ala^{144}	179.4	53.6	13.8			
Ala^{145}	177.4	51.8	15.6			
Thr^{146}	172.3	58.3	66.5	19.4		
Gly^{147}	172.2	42.6				
Met^{148}	169.7	53.6	27.4	35.0		15.6
Phe^{149}	170.5	54.0	39.0			129.1(C^{δ}),129.5(C^{ε}),127.3(C^{ζ})
Gly^{150}	168.6	41.5				
Lys^{151}	172.2	52.9	33.5	23.0	27.4	39.7
Ile^{152}	172.6	57.4	39.4	25.0	11.7	15.6
Val^{153}	173.0	59.6	31.1	18.3(0.87),17.9(0.95)		
Val^{154}	174.1	57.6	29.4	18.6(0.71),18.2(0.44)		
Lys^{155}	178.0	55.0	32.5	21.8	26.2	39.8

Numbers in parentheses refer to the methyl proton chemical shifts belonging to the respective methyl carbons of Val and Leu residues.

3.3. *Estimation of secondary structure from chemical shifts*

Carbon chemical shifts can provide a fairly reliable indicator of secondary structure [21]. Especially useful in this regard are the backbone chemical shifts, $^{13}C^{\alpha}$ and ^{13}CO, with the $^{13}C^{\beta}$ being a less reliable measure. While in principle the ^{15}N chemical shift should also provide secondary structure information, in practice the shift is highly sensitive to the hydrogen bonding state of the amide proton, and is therefore not a very reliable indicator of secondary structure. The most common method of extracting secondary structure information from chemical shift data is to compare the observed shifts with a set of 'random coil' chemical shifts, thus obtaining a measure of the perturbation of the shift by the presence of the residue in a paricular secondary structure. It is found [22] that the presence of helix causes the $^{13}C^{\alpha}$ and ^{13}CO resonances to shift downfield from the random coil position and the presence of β-structure causes an

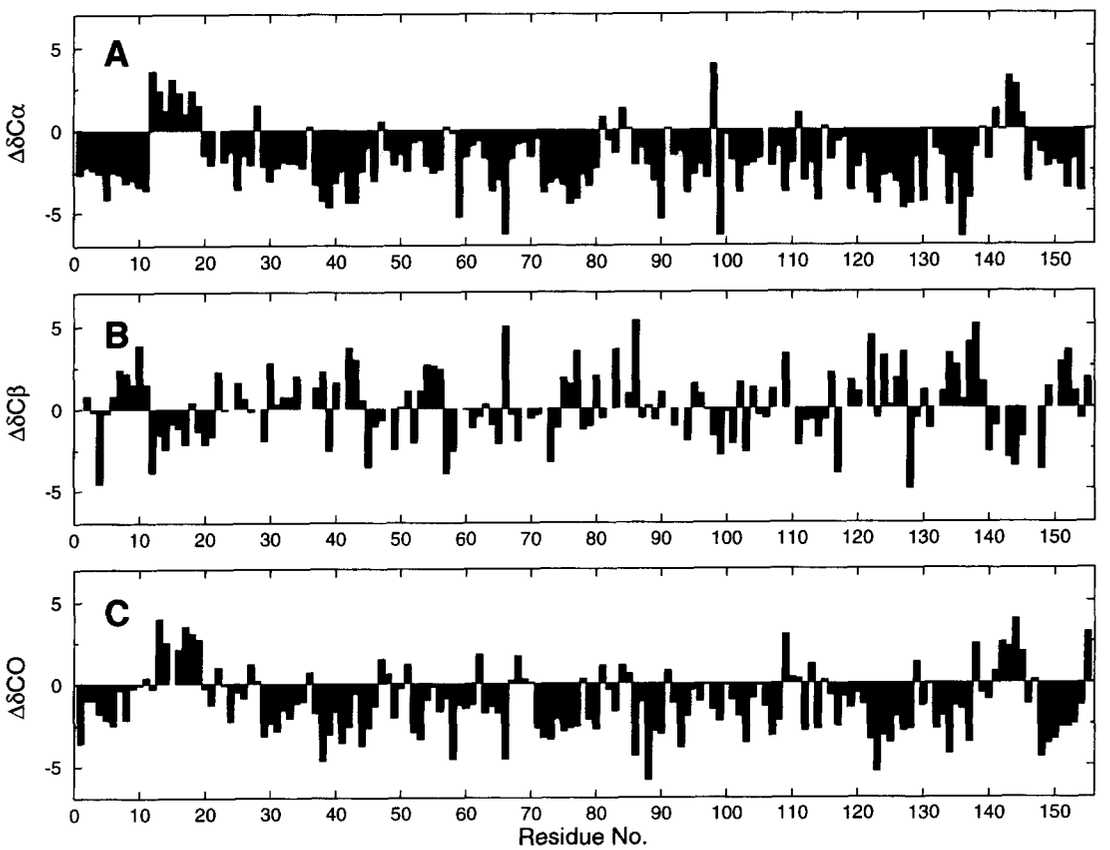

Fig. 4. Plot of the difference between the ^{13}C chemical shifts observed for rusticyanin and random-coil shifts for (A) $^{13}C^{\alpha}$ (B) $^{13}C^{\beta}$ (C) ^{13}CO.

Table 2
^{1}H resonance assignments for the proline residues in rusticyanin

Proline	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$
Pro^{13}	4.20	1.75,2.30	2.17,1.99	4.03,3.88
Pro^{47}	4.35	2.42,1.87	2.17,1.96	3.76,3.50
Pro^{50}	4.41	1.72,2.30	2.09,1.96	3.95,3.57
Pro^{52}	3.41	1.73,−0.60	1.52	3.17,3.30
Pro^{62}	4.66	1.73,1.80	2.21,2.17	3.71,3.52
Pro^{67}	4.76	1.78,2.30	2.26,2.00	3.15,3.95
Pro^{94}	2.31	1.40,0.95	1.45,1.65	2.01,2.60
Pro^{95}	3.90	2.20,1.85	1.87,1.77	3.28,3.38
Pro^{100}	4.21	1.70,2.00	2.05	4.06,3.90
Pro^{104}	4.49	1.93,2.30	1.97,2.15	3.61,3.55
Pro^{113}	4.27	1.53,2.13	1.63	2.27,1.76
Pro^{115}	4.37	1.53,1.92	1.63	3.67,3.92
Pro^{129}	3.71	−0.20,1.08	1.27,0.18	1.86,3.75
Pro^{141}	4.59	1.83,2.13	2.13,2.18	3.47,4.05

upfield shift. The effects on the C^{β} chemical shifts are much weaker and are reversed in sense. The difference between the observed and random coil shifts (obtained from [22]) for the C^{α}, C^{β} and CO chemical shifts is shown in Fig. 4. It is immediately obvious that the resonances of the majority of the molecule show an upfield shift in the measured $^{13}C^{\alpha}$ and ^{13}CO chemical shifts, indicative of β-structure. Two major regions are exceptions, between residues 11–20 and 138–143, which show a marked downfield shift, indicating the presence of two short stretches of helix. Several single residues throughout the chain show downfield shifts in both $^{13}C^{\alpha}$ and ^{13}CO, probably indicative of turns between β-strands. This estimation of the secondary structure of rusticyanin is entirely consistent with previous estimates obtained from NOE information and coupling constants [5].

These results indicate that the global fold of rusticyanin resembles those of other Type 1 copper proteins, and probably consists of a β-barrel or β-sandwich with the copper site at one end, as has been proposed [5]. Work is continuing on this system to calculate high-resolution solution structures of the protein in order to answer the most interesting questions concerning the protein structure in the vicinity of the copper site and how this relates to the stability of the site at the low pH optimum of rusticyanin.

Acknowledgements: We thank Dr. Peter Wright for continuing help and encouragement, Drs. John Cavanagh, John Chung, Li Xiang and Shohei Koide for helpful discussions regarding the setup of NMR experiments and Garry Gippert for helpful suggestions for data analysis. This work was supported by Grant GM48374 from the National Institutes of Health.

References

[1] Razzell, W.E. and Trussell, P.C. (1963) J. Bacteriol. 85, 595–603.
[2] Cox, J.C. and Boxer, D.H. (1978) Biochem. J. 174, 497–502.
[3] Blake II, R.C. and Shute, E.A. (1987) J. Biol. Chem. 262, 14983–14989.
[4] Guss, J.M., Harrowell, P.R., Murata, M., Norris, V.A. and Freeman, H.C. (1986) J. Mol. Biol. 192, 361–387.
[5] Hunt, A.H., Toy-Palmer, A., Assa-Munt, N., Cavanagh, J., Blake II, R.C. and Dyson, H.J. (1994) J. Mol. Biol. 244, 370–384.
[6] Casimiro, D.R., Toy-Palmer, A., Blake II, R.C. and Dyson, H.J. (1995) Biochemistry (in press).
[7] Vuister, G.W. and Bax, A. (1990) J. Magn. Reson. 98, 428–435.
[8] Live, D.H., Davis, D.G., Agosta, W.C. and Cowburn, D. (1984) J. Am. Chem. Soc. 106, 6104–6105.
[9] Ikura, M., Kay, L.E. and Bax, A. (1991) J. Biomol. NMR 1, 299–304.
[10] Bax, A., Clore, G.M. and Gronenborn, A.M. (1990) J. Magn. Reson. 88, 425–431.
[11] Grzesiek, S. and Bax, A. (1993) J. Biomol. NMR 3, 185–204.
[12] Grzesiek, S. and Bax, A. (1992) J. Am. Chem. Soc. 114, 6291–6293.
[13] Grzesiek, S., Anglister, J. and Bax, A. (1993) J. Magn. Reson. Series B 101, 114–119.
[14] Grzesiek, S. and Bax, A. (1992) J. Magn. Reson. 96, 432–440.
[15] Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) J. Magn. Reson. 88, 496–514.
[16] Yamazaki, T., Yoshida, M. and Nagayama, K. (1993) Biochemistry 32, 5656–5669.
[17] Oda, Y., Toshio, Y., Nagayama, K., Kanaya, S., Kuroda, Y. and Nakamura, H. (1994) Biochemistry 33, 5275–5284.
[18] Bax, A., Delaglio, F., Grzesiek, S. and Vuister, G.W. (1994) J. Biomol. NMR 4, 787–797.
[19] Grathwohl, C. and Wüthrich, K. (1976) Biopolymers 15, 2025–2041.
[20] Yao, J., Feher, V.A., Espejo, B.F., Reymond, M.T., Wright, P.E. and Dyson, H.J. (1994) J. Mol. Biol. 243, 736–753.
[21] Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) J. Mol. Biol. 222, 311–333.
[22] Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) J. Biomol. NMR 5, 67–81.